实验报告

成绩：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 班级学号 | 202015116 | 姓名 | 梁祺若 | 实验日期 | 2022-10-5 | 仪器编号 |  |
| 实验名称 | 实验3-5 R的决策树，人工神经网络 | | | | | | |

**1．实验题目**

R的决策树，人工神经网络

**2．实验步骤**

本题目由两部分组成：（1）利用R软件对专业棒球运动员数据和泰坦尼克号乘客信息做决策树，对Titanic(泰坦尼克号)上乘客的各类信息做分类树；（2）利用R软件对对neuralnet包数据做人工神经网络。

1. 利用R软件对专业棒球运动员数据和泰坦尼克号乘客信息做决策树，对Titanic(泰坦尼克号)上乘客的各类信息做分类树

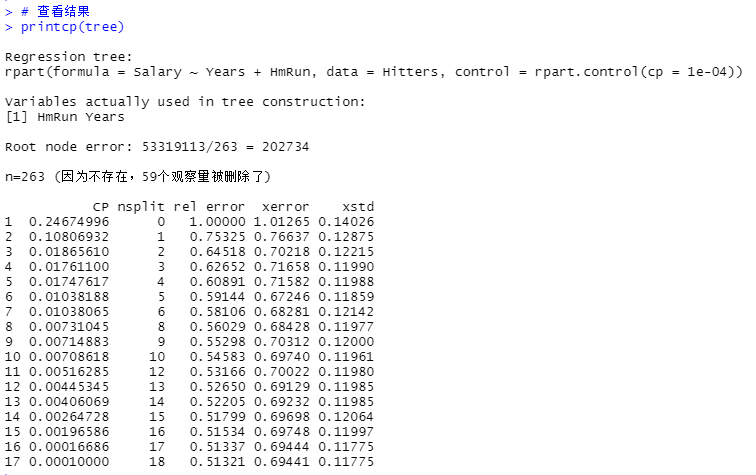
* 构建回归树（预测变量为home runs和years played，影响变量为运动员的Salary）
* 我们使用ISLR包中的Hitters数据集,它包括263个专业棒球运动员的各类信息。
* 导入所需包



* 构建初始回归树



* 查看结果



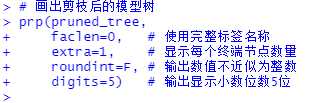
* 识别最佳CP值



* 基于最佳CP值对模型进行剪枝



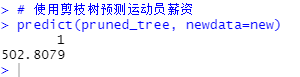
* 画出剪枝后的模型树



* 给定新的运动员信息

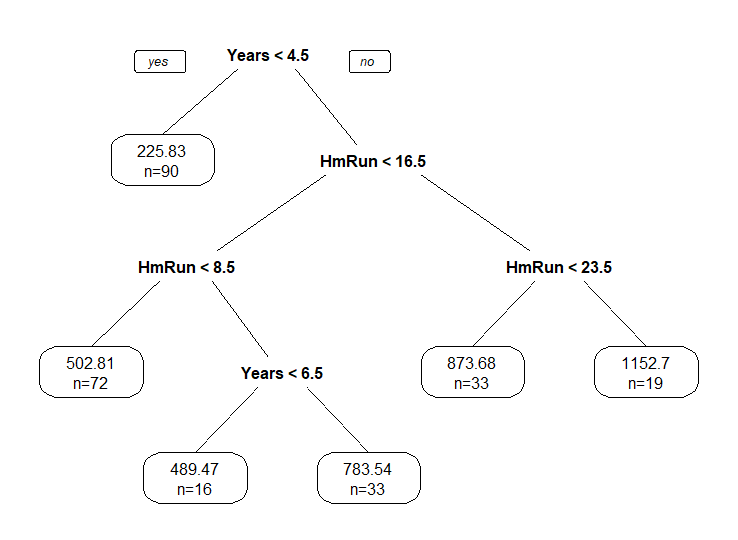


* 使用剪枝树预测运动员的薪资



* 查看结果





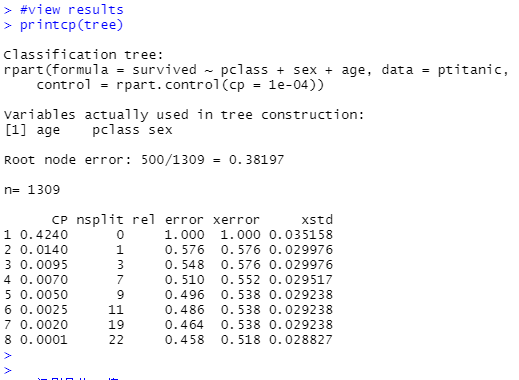
* 构建分类树（使用预测变量：pclass，sex和age，预测是否存活）
* 我们使用rpart.plot包中的ptitanic数据集, 它包含Titanic(泰坦尼克号)上乘客的各类信息。
* 导入所需包



* 构建分类树



* 查看结果



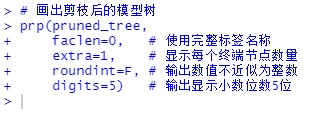
* 识别最佳CP值

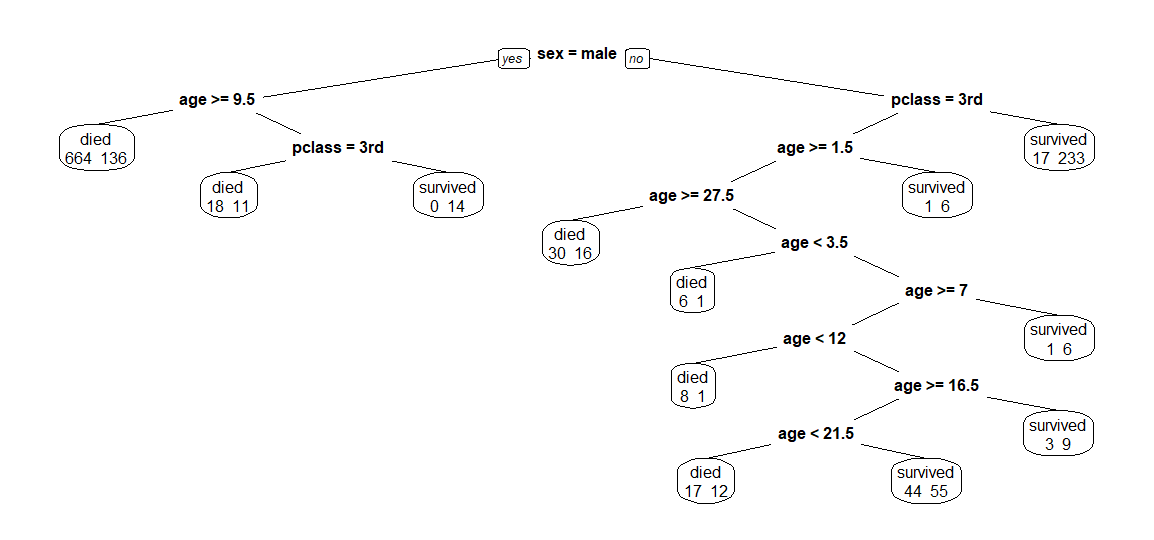


* 基于最佳CP值进行剪枝



* 画出剪枝后的模型树



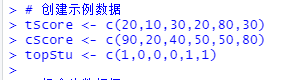


1. 利用R软件对neuralnet包数据做人工神经网络。

* 首先需要安装加载neuralnet包，然后创建示例数据。示例数据包括两类变量，特征变量和标签变量。特征变量包括知识技能和沟通交流两方面的得分，标签变量为二值标签（是否为尖子生）。
* 加载所需库



* 创建示例数据



* 组合为数据库



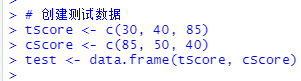
* 拟合神经网络模型



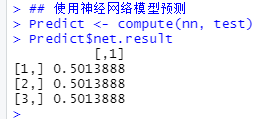
* 画出神经网络图



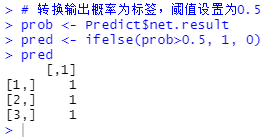
* 创建测试数据

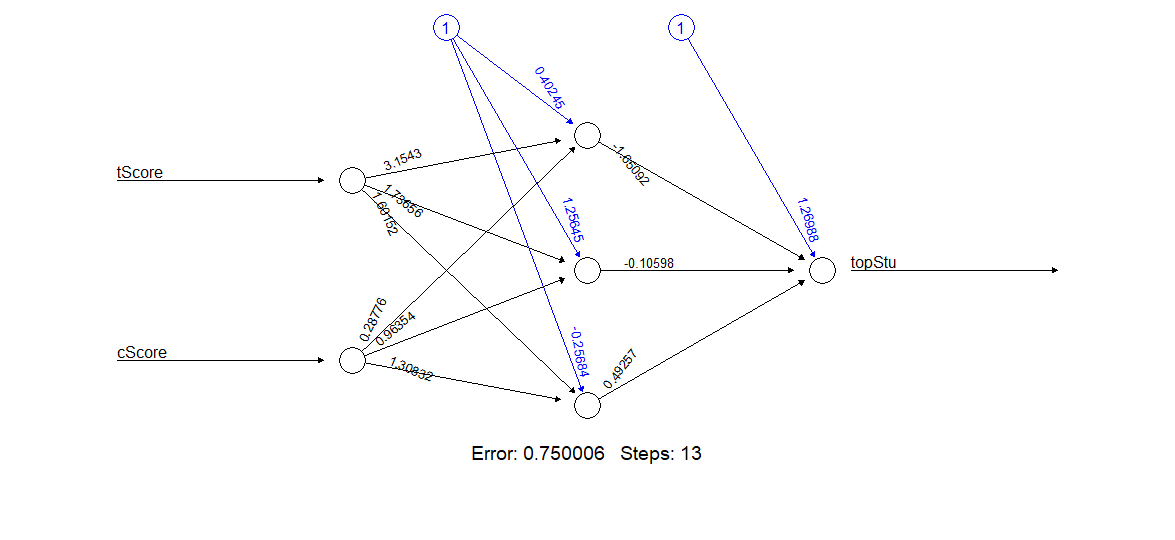


* 使用神经网络模型预测



* 转换输出概率为标签，阈值设置为0.5





**3．实验结果**

在本实验中，我们构建了两种不同的树，分别为回归树和分类树。并根据两种不同的树对不同的数据做个性化的分类处理，并做出预测。我们还运用了神经网络的算法对于数据进行处理。我们发现，神经网络更加灵活，不仅可以用于回归问题，也可以用于分类问题。